

01-24-00

A

01/21/00
jc760 U.S. PTO

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Our Docket No.: AMT-9704C
January 21, 2000

**CONTINUED PROSECUTION APPLICATION TRANSMITTAL
UNDER 37 CFR 1.53 (b)**

jc511 U.S. PTO
09/489254
01/21/00

Honorable Commissioner of
Patents and Trademarks
Box CPA
Washington, D.C. 20231

Dear Sir:

This is a request for filing a Continued Prosecution Application under 37 CFR 1.53(b) of application serial Number 08/932,078 filed May 27, 1997, entitled "Speaker Verification Method"

Enclosed are:

- ☒ 25 pages specification, 1 page of Abstract, and 7 sheet(s) of Drawings.
- ☒ The prior application is assigned of record to Ameritech Corporation.
- ☒ Copy of Declaration combined with Power of Attorney
- ☒ A preliminary amendment is enclosed to be entered in the new application after a filing date has been granted.
- ☐ An Information Disclosure Statement with Copies of references..
- ☒ The filing fee is calculated below:

CLAIMS AS FILED, LESS ANY CLAIMS CANCELLED BY AMENDMENT

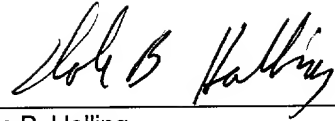
| | NUMBER FILED | NUMBER EXTRA | RATE | ADDITIONAL FEE |
|---------------------------|-----------------|-----------------|--------|-------------------|
| TOTAL CLAIMS | 10- 20 = | | X \$18 | = \$ 0.00 |
| INDEPENDENT CLAIMS | 0- 3 = | | X \$78 | = \$ 0.00 |
| MULTIPLE DEPENDENT CLAIMS | | | \$260 | = \$ 0.00 |
| BASIC FEE | | | | =\$690.00 |
| TOTAL FILING FEE | | | | =\$ 690.00 |

☐ Priority under 35 U.S.C 119 is claimed on the basis of prior application Serial No. , filed on . A certified copy of this foreign application was previously filed in the parent application.

☒ Amend the specification by inserting before the first line the sentence:
--This is a ☒ continuation division of copending US application Serial No. 08/863,444, filed on May 27, 1997, which application is hereby incorporated by reference herein."

Please Forward all Correspondence to

Law Office of Dale B. Halling
128 S. Tejon, Ste 202
Colorado Springs, CO 80903

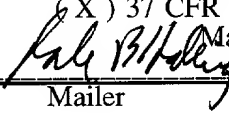


Dale B. Halling
Attorney/Agent for Applicants
Registration No. 38,170
Phone: (719) 447-1990
Fax: (719) 447-9815

CERTIFICATE OF MAILING

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being deposited on the date shown below with the United States Postal Service in an envelope addressed to the Commissioner of Patents and Trademarks, Washington, DC 20231,

() 37 CFR 1.8(a) as first class mail with sufficient postage
(X) 37 CFR 1.10 "Express Mail Post Office to Addressee"
Mailing Label No. EL367373170US



Mailer

Dale B. Halling

Printed Name of Mailer

Date: 1/21/00

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS: Bossemeyer, Jr. et al.

EXAMINER:

SERIAL NO.:

GROUP:

FILED:

CASE NO.: AMT-9704C(A00349C)

ENTITLED: Speaker Verification Method

Law Offices of Dale B. Halling
128 S. Tejon, Suite 202
Colorado Springs, CO 80903

January 21, 2000

Preliminary Amendment

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

Sir:

Please enter this amendment prior to examination.

In the specification:

Please change "'one)" to --"one")-- at page 9, line 10.

In The Claims:

Please cancel claims 1-21 and add new claims 22-31

22(new). A speaker verification method comprising the steps of:

_____ (a) generating a code book covering a plurality of speakers having a plurality of training utterances for each of the plurality of speakers, wherein a first utterance of the plurality of training utterances is not the same as a second utterance of the plurality of training utterances;

_____ (b) receiving a plurality of test utterances from one of the plurality of speakers, wherein a first of the plurality of test utterances is not the same as a second of the plurality of test utterances;

_____ (c) comparing each of the plurality of test utterances to each of the plurality of training utterances for the speaker to form a plurality of preliminary decisions, one preliminary decision of the plurality of preliminary decisions for each of the plurality of test utterances; and

_____ (d) combining the plurality of preliminary decisions to form a verification decision.

_____ 23(new). The method of claim 22, wherein step (d) further includes the steps of:

_____ (d1) weighting each of the plurality of preliminary decisions.

24(new). The method of claim 23, wherein the step of weighting includes determining a historical probability of false alarm for each of the plurality of utterances.

25(new). The method of claim 23, wherein the step of weighting includes evaluating a quality of the preliminary decision for each of the plurality of decisions.

26(new). The method of claim 22, wherein step (a) further includes the steps of:

(a1) separating the plurality of speakers into a male group and a female group;

(a2) determining a male variance vector from the male group;

(a3) determining a female variance vector from the female group.

27(new) The method of claim 26, wherein step (b) further includes the step of:

(b1) determining if the speaker is a male.

28(new). The method of claim 27, wherein step (c) further includes the steps of:

(c1) when the speaker is male using the male variance vector to determine a weighted Euclidean distance between each of the plurality of test utterances and the plurality of training utterances.

29(new) A method of speaker verification, comprising the steps:

(a) receiving a plurality of test utterances from a speaker;

(b) determining if the speaker is a male;

(c) when the speaker is male, using a male variance vector to determine a weighted Euclidean distance between each of the plurality of test utterances and the plurality of training utterances;

(d) forming a preliminary decision for each of the plurality of test utterances to form a plurality of preliminary decisions;

(e) combining the plurality of preliminary decisions to form a verification decision.

30(new). The method of claim 29, further including the steps of:

(f) when the speaker is not male, using a female variance vector to determine a weighted Euclidean distance between each of the plurality of test utterances and the plurality of training utterances;

(d) forming a preliminary decision for each of the plurality of test utterances to form a plurality of preliminary decisions;

(e) combining the plurality of preliminary decisions to form a verification decision.

31(new). A speaker verification method comprising the steps of:

- (a) receiving a plurality of different test utterances from a speaker;
- (b) comparing the plurality of different test utterances to a plurality of training utterance for the speaker to form a plurality of preliminary decisions;
- (c) combining the plurality of preliminary decisions to form a verification decision.

Remarks

The applicants have canceled claims 1-21 and have added new claims 22-31, which are fully supported by the specification.

Respectfully submitted,

(Bossemeyer, Jr. et al.)

By 

Attorney for the Applicant

Dale B. Halling

Phone: (719) 447-1990

Fax: (719) 447-0983

I hereby certify that an Amendment is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on:

1/21/00

Date



Signature (Dale B. Halling)

SPEAKER VERIFICATION METHOD

Field of the Invention

The present invention is related to the field of speech recognition systems and more particularly to a speaker verification method.

Background of the Invention

Speech recognition and speaker verification use similar analysis tools to achieve its goals. An input utterance is first processed to determine its essential characteristics. Typically, input utterances are converted to cepstrum coefficients. A cepstrum is an inverse Fourier transform of the log power spectrum. In a training phase the cepstrum coefficients are saved to form code books for specific utterances. For instance, a code book might have codes for the numeral zero through nine. In speech recognition, an input utterance is compared to the codes (training utterances) in the code book to determine which is most similar. In speech recognition the code is a generalized representation of many people's way of forming an utterance (e.g., "zero"). In speaker verification the codes represent the individual characteristics of the

speaker and the verification system tries determine if a person's code is more similar to an input utterance, than an impostor code. As a result the codes in a speaker verification system emphasis individual characteristics, while in a speech recognition system the codes generalize over many individual speakers. Speaker verification has potential applications in a number of voice activated systems, such as banking over the telephone. Unfortunately, present speaker verification systems have not proven reliable enough for these applications.

Thus there exists a need for an improved speaker verification system.

Summary of the Invention

An improved speaker verification method consist of the following steps: (1) generating a code book covering a number of speakers having a number of training utterances for each of the speakers; (2) receiving a number of test utterances from a speaker; (3) comparing each of the test utterances to each of the training utterances for the speaker to form a number of decisions, one decision for each of the number of test utterances, (4) weighting each of the decisions to form a number of weighted decisions; and (5) combining the plurality of weighted decisions to form a verification decision.

Brief Description of the Drawings

FIG. 1 is a block diagram of an embodiment of a speaker verification system;

5 FIG. 2 is a flow chart of an embodiment of the steps used to form a speaker verification decision;

FIG. 3 is a flow chart of an embodiment of the steps used to form a code book for a speaker verification decision;

10 FIG. 4 is a flow chart of an embodiment of the steps used to form a speaker verification decision;

FIG. 5 is a schematic diagram of a dial-up service that incorporates a speaker verification method;

FIG. 6 is a flow chart of an embodiment of the steps used in a dial-up service; and

15 FIG. 7 is a flow chart of an embodiment of the steps used in a dial-up service.

Detailed Description of the Drawings

Several improvements in speaker verification methods are described and then a dial-up service that can incorporate these improvements is explained. FIG. 1 is a block diagram of an embodiment of a speaker verification system 10. It is important to note that the speaker verification system can be physically implemented in a number of ways. For instance, the system can be implemented as software in a general purpose computer connected to a microphone; or the system can be implemented as firmware in a general purpose microprocessor connected to memory and a microphone; or the system can be implemented using a Digital Signal Processor (DSP), a controller, a memory, and a microphone controlled by the appropriate software. Note that since the process can be performed using software in a computer, then a computer readable storage medium containing computer readable instructions can be used to implement the speaker verification method. These various system architectures are apparent to those skilled in the art and the particular system architecture selected will depend on the application.

A microphone 12 receives an input speech and converts the sound waves to an electrical signal. A feature extractor 14 analyzes the electrical signal and extracts key features of the speech. For instance, the feature extractor first digitizes the electrical signal. A cepstrum of the digitized signal is then performed to determine the cepstrum coefficients. In another embodiment, a linear predictive analysis is used

to find the linear predictive coding (LPC) coefficients. Other feature extraction techniques are also possible.

A switch 16 is shown attached to the feature extractor 14. This switch 16 represents that a different path is used in the training phase than in the verification phase. In the training phase the cepstrum coefficients are analyzed by a code book generator 18. The output of the code book generator 18 is stored in the code book 20. In one embodiment, the code book generator 18 compares samples of the same utterance from the same speaker to form a generalized representation of the utterance for that person. This generalized representation is a training utterance in the code book. The training utterance represents the generalized cepstrum coefficients of a user speaking the number "one" as an example. A training utterance could also be a part of speech, a phoneme, or a number like "twenty one" or any other segment of speech. In addition to the registered users samples, utterances are taken from a group of non-users. These utterances are used to form a composite that represents an impostor code having a plurality of impostor utterances.

In one embodiment, the code book generator 18 determines whether the speaker (users and non-users) is male or female. The male training utterances (male group) are aggregated to determining a male variance vector. The female training utterances (female group) are aggregated to determine a female variance vector. These gender specific variance vectors will be used when calculating a weighted Euclidean distance (measure of closeness) in the verification phase.

In the verification phase the switch 16 connects the feature extractor 14 to the comparator 22. The comparator 22 performs a mathematical analysis of the closeness between a test utterance from a speaker with a training utterance stored in the code book 20 and
5 between the test utterance and an impostor utterance. In one embodiment, a test utterance such as a spoken "one" is compared with the "one" training utterance and the "one" impostor utterance in the code book 20. The comparator 22 determines a measure of closeness between the "one" training utterance, the "one" test utterance and the
10 "one" impostor utterance. When the test utterance is closer to the training utterance than the impostor utterance, the speaker is verified as the true speaker. Otherwise the speaker is determined to be an impostor. In one embodiment, the measure of closeness is a modified weighted Euclidean distance. The modification in one embodiment
15 involves using a generalized variance vector instead of an individual variance vector for each of the registered users. In another embodiment, a male variance vector is used for male speakers and a female variance vector is used for a female speaker.

A decision weighting and combining system 24 uses the measure
20 of closeness to determine if the test utterance is closest to the training utterance or the impostor utterance. When the test utterance is closer to the training utterance than impostor utterance, a verified decision is made. When the test utterance is not closer to the training utterance than the impostor utterance, an un-verified decision is made. These are
25 preliminary decisions. Usually, the speaker is required to speak several utterances (e.g., "one", "three", "five", "twenty one"). A decision is made

for each of these test utterances. Each of the plurality of decisions is weighted and combined to form the verification decision.

The decisions are weighted because not all utterances provide equal reliability. For instance, "one" could provide a much more reliable decision than "eight". As a result, a more accurate verification decision can be formed by first weighting the decisions based on the underlying utterance. Two weighting methods can be used. One weighting method uses a historical approach. Sample utterances are compared to the training utterances to determine a probability of false alarm P_{FA} (speaker is not impostor but the decision is impostor) and a probability of miss P_M (speaker is impostor but the decision is true speaker). The P_{FA} and P_M are probability of errors. These probability of errors are used to weight each decision. In one embodiment the weighting factors (weight) are described by the equation below:

$$a_i = \log \frac{1 - P_M}{P_{FAi}} \quad \text{Decision is Verified (True Speaker)}$$

$$a_i = \log \frac{P_M}{1 - P_{FAi}} \quad \text{Decision is Not Verified (Impostor)}$$

When the sum of the weighted decisions is greater than zero, then the verification decision is a true speaker. Otherwise the verification decision is an impostor.

The other method of weighting the decisions is based on an immediate evaluation of the quality of the decision. In one embodiment, this is calculated by using a Chi-Squared detector. The decisions are then weighted on the confidence determined by the Chi-Squared
5 detector. In another embodiment, a large sample approximation is used. Thus if the test statistics are t , find b such that $c^2(b) = t$. Then a decision is an impostor if it exceeds the $1-a$ quantile of the c^2 distribution.

One weighting scheme is shown below:

10

$$\begin{aligned} &1.5, \text{ if } b > c_{\text{accept}} \\ &1.0, \text{ if } 1-a \leq b \leq c_{\text{accept}} \\ &-1.0, \text{ if } c_{\text{reject}} \leq b \leq 1-a \\ &-1.25, \text{ if } b < c_{\text{reject}} \end{aligned}$$

15

When the sum of the weighted decisions is greater than zero, then the verification decision is a true speaker. When the sum of the weighted decision is less than or equal to zero, the decision is an impostor.

20

In another embodiment, the feature extractor 14 segments the speech signal into voiced sounds and unvoiced sounds. Voiced sounds generally include vowels, while most other sounds are unvoiced. The unvoiced sounds are discarded before the cepstrum coefficients are calculated in both the training phase and the verification phase.

These techniques of weighting the decisions, using gender dependent cepstrums and only using voiced sounds can be combined or used separately in a speaker verification system.

FIG. 2 is a flow chart of an embodiment of the steps used to form a speaker verification decision. The process starts, at step 40, by
5 generating a code book at step 42. The code book has a plurality of training utterances for each the plurality of speakers (registered users, plurality of people) and a plurality of impostor utterances. The training utterances in one embodiment are the cepstrum coefficients for a
10 particular user speaking a particular utterance (e.g., "one). The training utterances are generated by a user speaking the utterances. The cepstrum coefficients of each of the utterances are determined to form the training utterances. In one embodiment a speaker is asked to repeat the utterance and a generalization of the two utterances is saved as the
15 training utterance. In another embodiment both utterances are saved as training utterances.

In one embodiment, a data base of male speakers is used to determine a male variance vector and a data base of female speakers is used to determine a female variance vector. In another embodiment,
20 the data bases of male and female speakers are used to form a male impostor code book and a female impostor code book. The gender specific variance vectors are stored in the code book. At step 44, a plurality of test utterances (input set of utterances) from a speaker are received. In one embodiment the cepstrum coefficients of the test
25 utterances are calculated. Each of the plurality of test utterances are compared to the plurality of training utterances for the speaker at step

46. Based on the comparison, a plurality of decision are formed, one for each of the plurality of training utterances. In one embodiment, the comparison is determined by a Euclidean weighted distance between the test utterance and the training utterance and between the test utterance and an impostor utterance. In another embodiment, the Euclidean weighted distance is calculated with the male variance vector if the speaker is a male or the female variance vector if the speaker is a female. Each of the plurality of decisions are weighted to form a plurality of weighted decisions at step 48. The weighting can be based on historical error rates for the utterance or based on a confidence level (confidence measure) of the decision for the utterance. The plurality of weighted decisions are combined at step 50. In one embodiment, at the step of combining involves summing the weighted decisions. A verification decision is then made based on the combined weighted decisions at step 52, ending the process at step 54. In one embodiment if the sum is greater than zero, the verification decision is the speaker is a true speaker, otherwise the speaker is an impostor.

FIG. 3 is a flow chart of an embodiment of the steps used to form a code book for a speaker verification decision. The process starts, at step 70, by receiving an input utterance at step 72. In one embodiment, the input utterances are then segmented into a voiced sounds and an unvoiced sounds at step 74. The cepstrum coefficients are then calculated using the voiced sounds at step 76. The coefficients are stored as a training utterance for the speaker at step 78. The process then returns to step 72 for the next input utterance, until all the training utterances have been stored in the code book.

FIG. 4 is a flow chart of an embodiment of the steps used to form a speaker verification decision. The process starts, at step 100, by receiving input utterances at step 102. Next, it is determined if the speaker is male or female at step 104. In a speaker verification application, the speaker purports to be someone in particular. If the person purports to be someone that is a male, then the speaker is assumed to be male even if the speaker is a female. The input utterances are then segmented into a voiced sounds and an unvoiced sounds at step 106. Features (e.g., cepstrum coefficients) are extracted from the voiced sounds to form the test utterances, at step 108. At step 110, the weighted Euclidean distance (WED) is calculated using either a generalized male variance vector if the purported speaker is a male. When the purported speaker is a female, the female variance vector is used. The WED is calculated between the test utterance and the training utterance for the speaker and the test utterance and the male (or female if appropriate) impostor utterance. A decision is formed for each test utterance based on the WED at step 112. The decisions are then weighted based on a confidence level (measure of confidence) determined using a Chi-squared detector at step 114. The weighted decisions are summed at step 116. A verification decision is made based on the sum of the weighted decisions at step 118.

Using the speaker verification decisions discussed above results in an improved speaker verification system, that is more reliable than present techniques.

A dial-up service that uses a speaker verification method as described above is shown in FIG. 5. The dial-up service is shown as a

banking service. A user dials a service number on their telephone 150. The public switched telephone network (PSTN) 152 then connects the user's phone 150 with a dial-up service computer 154 at a bank 156. The dial-up service need not be located within a bank. The service will be explained in conjunction with the flow chart shown in FIG. 6. The process starts, at step 170, by dialing a service number (communication service address, number) at step 172. The user (requester) is then prompted by the computer 154 to speak a plurality of digits (access code, plurality of numbers, access number) to form a first utterance at step 174. The digits are recognized using speaker independent voice recognition at step 176. When the user has used the dial-up service previously, verifying the user based on the first utterance at step 178. When the user is verified as a true speaker at step 178, allowing access to the dial-up service at step 180. When the user cannot be verified, requesting the user input a personal identification number (PIN) at step 182. The PIN can be entered by the user either by speaking the PIN or by enter the PIN on a keypad. At step 184 it is determined if the PIN is valid. When the PIN is not valid, the user is denied access at step 186. When the PIN is valid the user is allowed access to the service at step 180. Using the above method the dial-up service uses a speaker verification system as a PIN option, but does not deny access to the user if it cannot verify the user.

FIG. 7 is a flow chart of another embodiment of the steps used in a dial-up service. The process starts, step 200, by the user speaking an access code to form a plurality of utterances at step 202. At step 204 it is determined if the user has previously accessed the service. When the

user has previously used the service, the speaker verification system attempts to verify the user (identity) at step 206. When the speaker verification system can verify the user, the user is allowed access to the system at step 208. When the system cannot verify the user, a PIN is requested at step 210. Note the user can either speak the PIN or enter the PIN on a keypad. At step 212 it is determined if the PIN is valid. When the PIN is not valid the user is denied access at step 214. When the PIN is valid, the user is allowed access at step 208.

When the user has not previously accessed the communication service at step 204, the user is requested to enter a PIN at step 216. At step 218 it is determined if the PIN is valid at step 218. When the PIN is not valid, denying access to the service at step 220. When the PIN is valid the user is asked to speak the access code a second time to form a second utterance (plurality of second utterances) at step 222. The similarity between the first utterance (step 202) and the second utterance is compared to a threshold at step 224. In one embodiment the similarity is calculated using a weighted Euclidean distance. When the similarity is less than or equal to the threshold, the user is asked to speak the access code again at step 222. In this case the second and third utterances would be compared for the required similarity. In practice, the user would not be required to repeat the access code at step 222 more than once or twice and the system would then allow the user access. When the similarity is greater than the threshold, storing a combination of the two utterances as at step 226. In another embodiment both utterances are stored as reference utterances. Next access to the service is allowed at step 208. The reference utterance

(plurality of reference utterances, reference voiced sounds) is used to verify the user the next time they access the service. Note that the speaker verification part of the access to the dial-up service in one embodiment uses all the techniques discussed for a verification process.

5 In another embodiment the verification process only uses one of the speaker verification techniques. Finally, in another embodiment the access number has a predetermined digit that is selected from a first set of digits (predefined set of digits) if the user is a male. When the user is a female, the predetermined digit is selected from a second set of digits.

10 This allows the system to determine if the user is suppose to be a male or a female. Based on this information, the male variance vector or female variance vector is used in the speaker verification process.

Thus there has been described an improved speaker verification method and a service that takes advantage of the speaker verification method. While the invention has been described in conjunction with

15 specific embodiments thereof, it is evident that many alterations, modifications, and variations will be apparent to those skilled in the art in light of the foregoing description. Accordingly, it is intended to embrace all such alterations, modifications, and variations in the

20 appended claims.

Claims

What is claimed is:

5

1. A speaker verification method comprising the steps of:

10

(a) generating a code book covering a plurality of speakers having a plurality of training utterances for each of the plurality of speakers;

(b) receiving a plurality of test utterances from a speaker;

15

(c) comparing each of the plurality of test utterances to each of the plurality of training utterances for the speaker to form a plurality of decisions, one decision of the plurality of decisions for each of the plurality of test utterances;

(d) weighting each of the plurality of decisions to form a plurality of weighted decisions; and

(e) combining the plurality of weighted decisions to form a verification decision.

20

2. The method of claim 1, wherein step (c) further includes the step of:

25

(c1) comparing each of the plurality of test utterances to each of a plurality of impostor utterances.

3. The method of claim 1, wherein step (d) further includes the steps of:

5 (d1) determining a measure of confidence for each of the plurality of decisions;

(d2) assigning a weight for each of the plurality of decisions based on the measure of confidence.

10 4. The method of claim 1, wherein step (a) further includes the steps of:

(a1) separating the plurality of speakers into a male group and a female group;

15 (a2) determining a male variance vector from the male group;

(a3) determining a female variance vector from the female group.

5. The method of claim 1, wherein step (c) further including the steps of:

(c1) determining if the speaker of the plurality of test utterances is a male or a female;

(c2) when the speaker is male, using the male variance vector to determine a weighted Euclidean distance between each of the plurality of test utterances and each of the plurality of training utterances for the speaker;

(c3) forming a decision for each of the plurality of test utterances based on the weighted Euclidean distance.

6. The method of claim 4, wherein step (c) further including the steps of:

(c1) determining if the speaker of the test utterances is a male or a female;

(c2) when the speaker is female, using the female variance vector to determine a weighted Euclidean distance between each of the plurality of test utterances and each of the plurality of training utterances for the speaker;

(c3) forming a decision for each of the plurality of test utterances based on the weighted Euclidean distance.

7. The method of claim 1, wherein step (a) further includes the steps of:

(a1) receiving a sample utterance;

5 (a2) segmenting the sample utterance into a voiced sounds and an unvoiced sounds;

(a3) storing the voiced sounds as one of the plurality of training utterances.

10 8. The method of claim 7, wherein step (b) further includes the steps of:

(b1) receiving an input set of utterances;

15 (b2) segmenting the input set of utterances into the voiced sounds and the unvoiced sounds;

(b3) storing the voiced sounds to form the plurality of test utterances.

9. A method of speaker verification, comprising the steps:

(a) generating a code book containing a plurality of training utterances for a plurality of people and a male variance vector, a female variance vector and a plurality of impostor utterances;

(b) receiving a plurality of test utterances from a speaker;

(c) determining if the speaker is a male;

(d) when the speaker is male, using the male variance vector to determine a weighted Euclidean distance between each of the plurality of test utterances and the plurality of training utterances;

(e) determining a weighted Euclidean distance between each of the plurality of test utterances and the plurality of impostor utterances;

(f) forming a decision for each of the plurality of test utterances to form a plurality of decisions; and

(g) combining the plurality of decisions to form the verification decision.

10. The method of claim 9, wherein step (f) further includes the step of:

(f1) comparing the weighted Euclidean distance for each
5 the plurality of training utterances to the weighted Euclidean distance for each of the plurality of impostor utterances to form a comparison;

(f2) forming a decision based on the comparison.

10 11. The method of claim 9, wherein step (g) further includes the steps of:

(g1) weighting each of the plurality of decisions based on a confidence measure to form a plurality of weighted decisions;

15 (g2) summing the plurality of weighted decisions to form a verification decision.

12. The method of claim 9, wherein step (a) further includes the steps of:

20

(a1) receiving a sample utterance;

(a2) segmenting the sample utterance into a voiced sounds and an unvoiced sounds;

25 (a3) storing the voiced sounds as one of the plurality of training utterances.

13. The method of claim 12, wherein step (b) further includes the steps of:

- 5 (b1) receiving an input set of utterances;
- (b2) segmenting the input set of utterances into the voiced sounds and the unvoiced sounds;
- (b3) storing the voiced sounds to form the plurality of test utterances.

10

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211

14. A computer readable storage medium containing computer readable instructions that when executed by a computer performs the following steps:

5 (a) generates a code book covering a plurality of speakers having a plurality of training utterances for each of the plurality of speakers;

(b) receives a plurality of test utterances from a speaker;

10 (c) compares each of the plurality of test utterances to each of the plurality of training utterances in the code book to form a plurality of decisions, one decision of the plurality of decisions for each of the plurality of test utterances;

(d) weights each of the plurality of decisions to form a plurality of weighted decisions; and

15 (e) combines the plurality of weighted decisions to form a verification decision.

15. The method of claim 14, wherein step (c) further includes the step of:

20

(c1) comparing each of the plurality of test utterances to each of a plurality of impostor utterances.

16. The computer readable storage medium of claim 14,
wherein step (d) further includes the steps of:

(d1) determines a measure of confidence for each of the
5 plurality of decisions;

(d2) assigns a weight for each of the plurality of decisions
based on the measure of confidence.

17. The computer readable storage medium of claim 14,
10 wherein step (a) further includes the steps of:

(a1) separates the plurality of speakers into a male group
and a female group;

(a2) determines a male variance vector from the male
15 group;

(a3) determines a female variance vector from the female
group.

18. The computer readable storage medium of claim 14,
wherein step (c) further including the steps of:

5 (c1) determines if the speaker of the plurality of test
utterances is a male or a female;

(c2) when the speaker is male, using the male variance
vector to determine a weighted Euclidean distance between each of
the plurality of test utterances and each of the plurality of training
utterances for the speaker;

10 (c3) forms a decision for each of the plurality of test
utterances based on the weighted Euclidean distance.

19. The computer readable storage medium of claim 17,
wherein step (c) further including the steps of:

15 (c1) determines if the speaker of the plurality of test
utterances is a male or a female;

(c2) when the speaker is female, using the female
variance vector to determine a weighted Euclidean distance between
20 each of the plurality of test utterances and each of the plurality of
training utterances for the speaker;

(c3) forms a decision for each of the plurality of test
utterances based on the weighted Euclidean distance.

20. The computer readable storage medium of claim 14,
wherein step (a) further includes the steps of:

(a1) receives a sample utterance;

5 (a2) segments the sample utterance into a voiced sounds
and an unvoiced sounds;

(a3) stores the voiced sounds as one of the plurality of
training utterances.

10 21. The computer readable storage medium of claim 20,
wherein step (b) further includes the steps of:

(b1) receives an input set of utterances;

15 (b2) segments the input set of utterances into the voiced
sounds and the unvoiced sounds;

(b3) stores the voiced sounds to form the plurality of test
utterances.

SPEAKER VERIFICATION METHOD

5

Abstract of the Disclosure

A speaker verification method consist of the following steps:

- 10 (1) generating a code book (42) covering a number of speakers having a number of training utterances for each of the speakers; (2) receiving a number of test utterances (44) from a speaker; (3) comparing (46) each of the test utterances to each of the training utterances for the speaker to form a number of decisions, one
- 15 decision for each of the number of test utterances; (4) weighting each of the decisions (48) to form a number of weighted decisions; and (5) combining (50) the plurality of weighted decisions to form a verification decision (52).

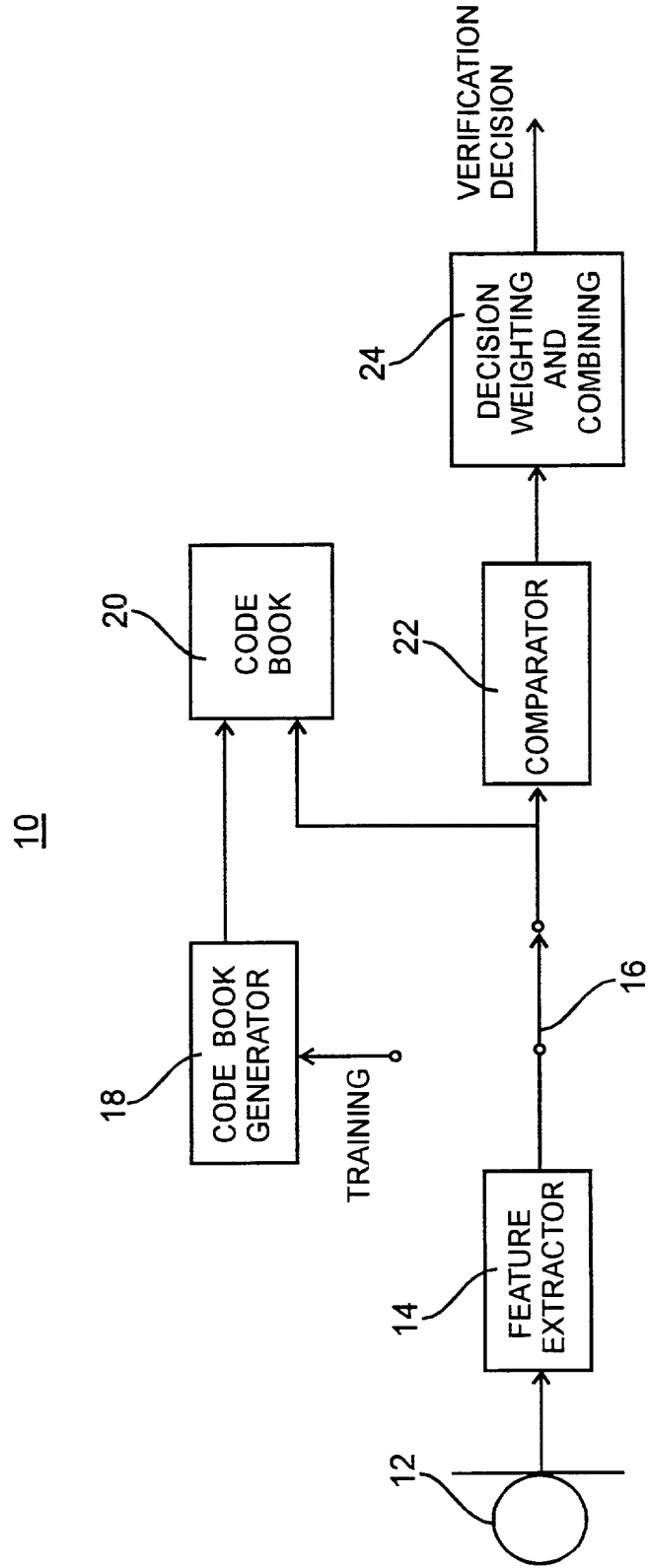
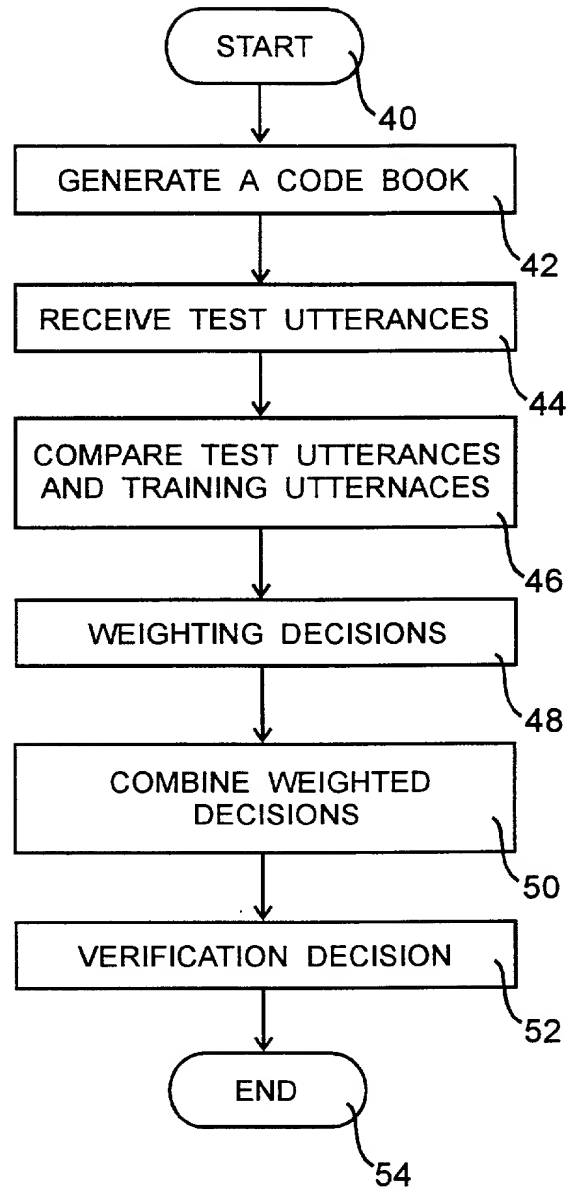
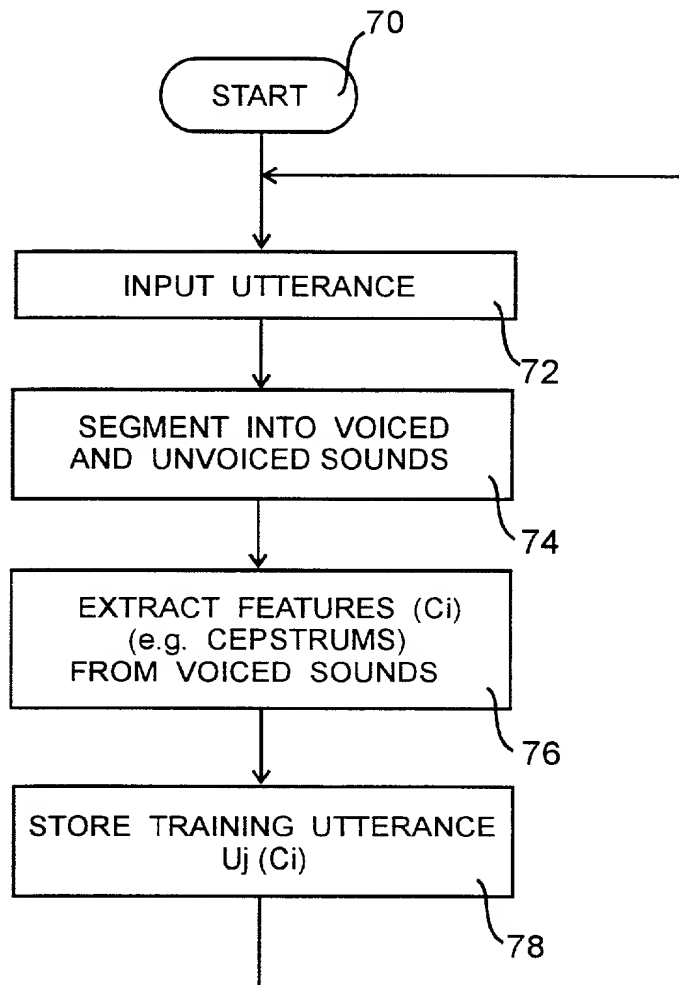
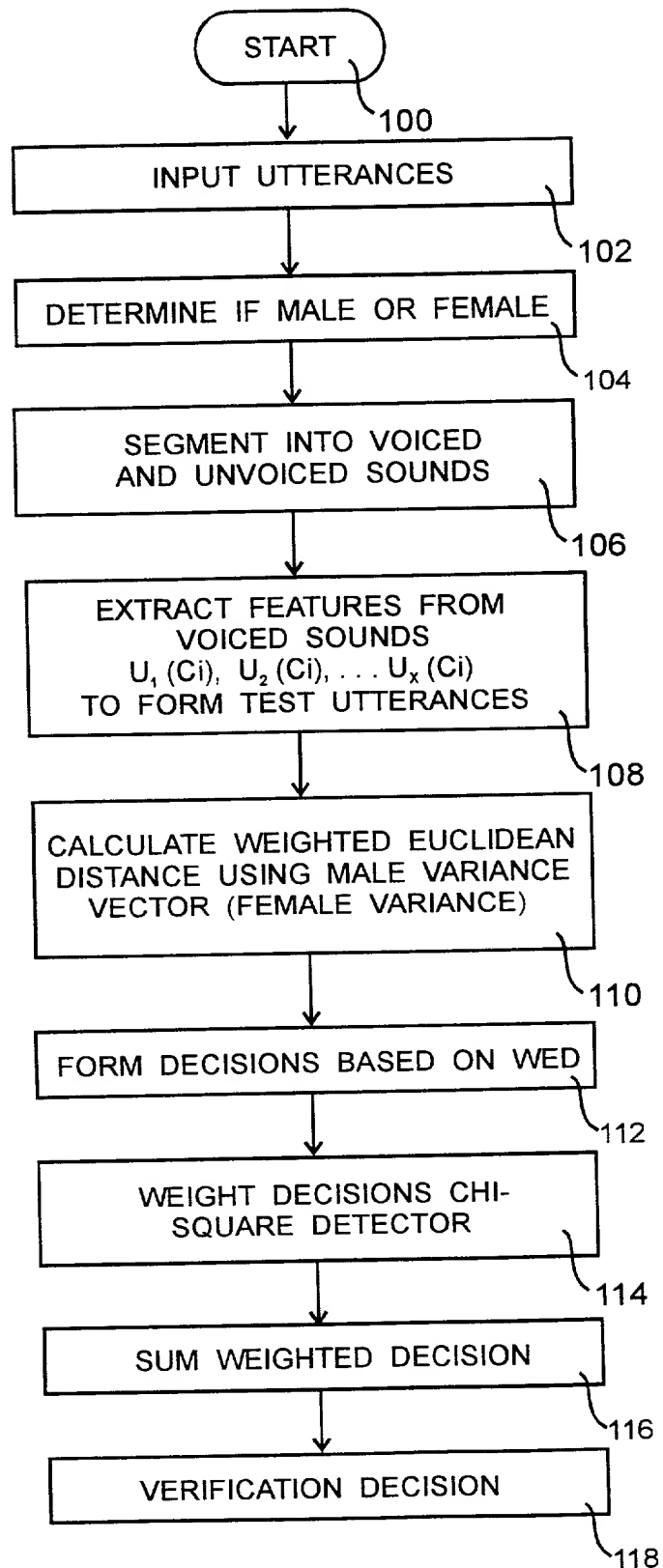


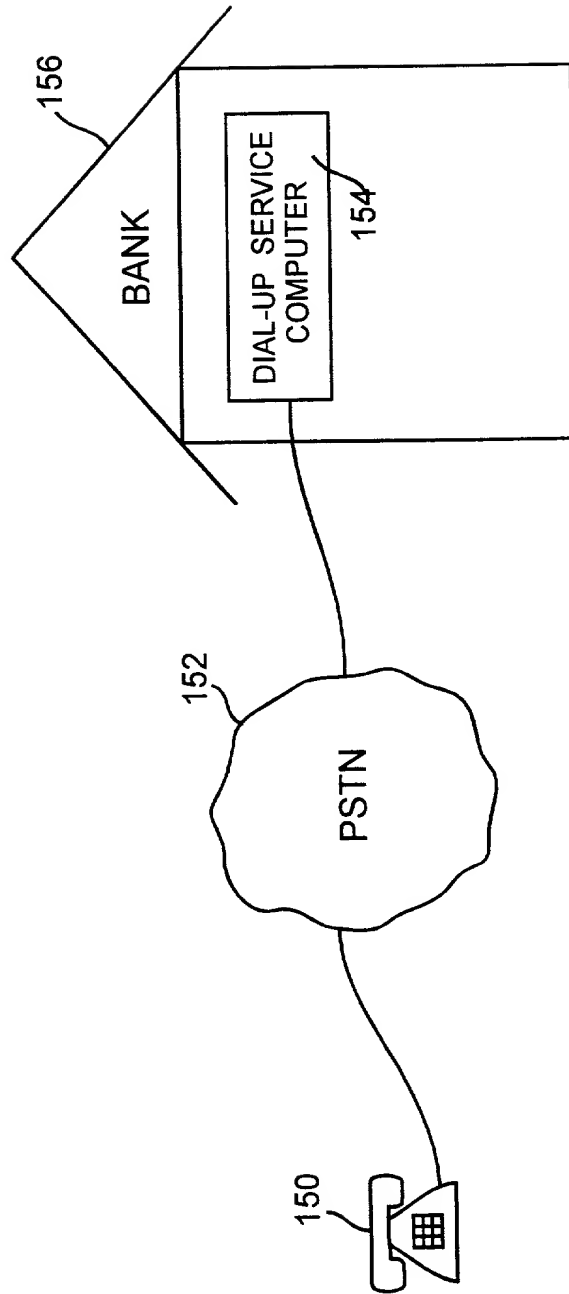
Fig. 1

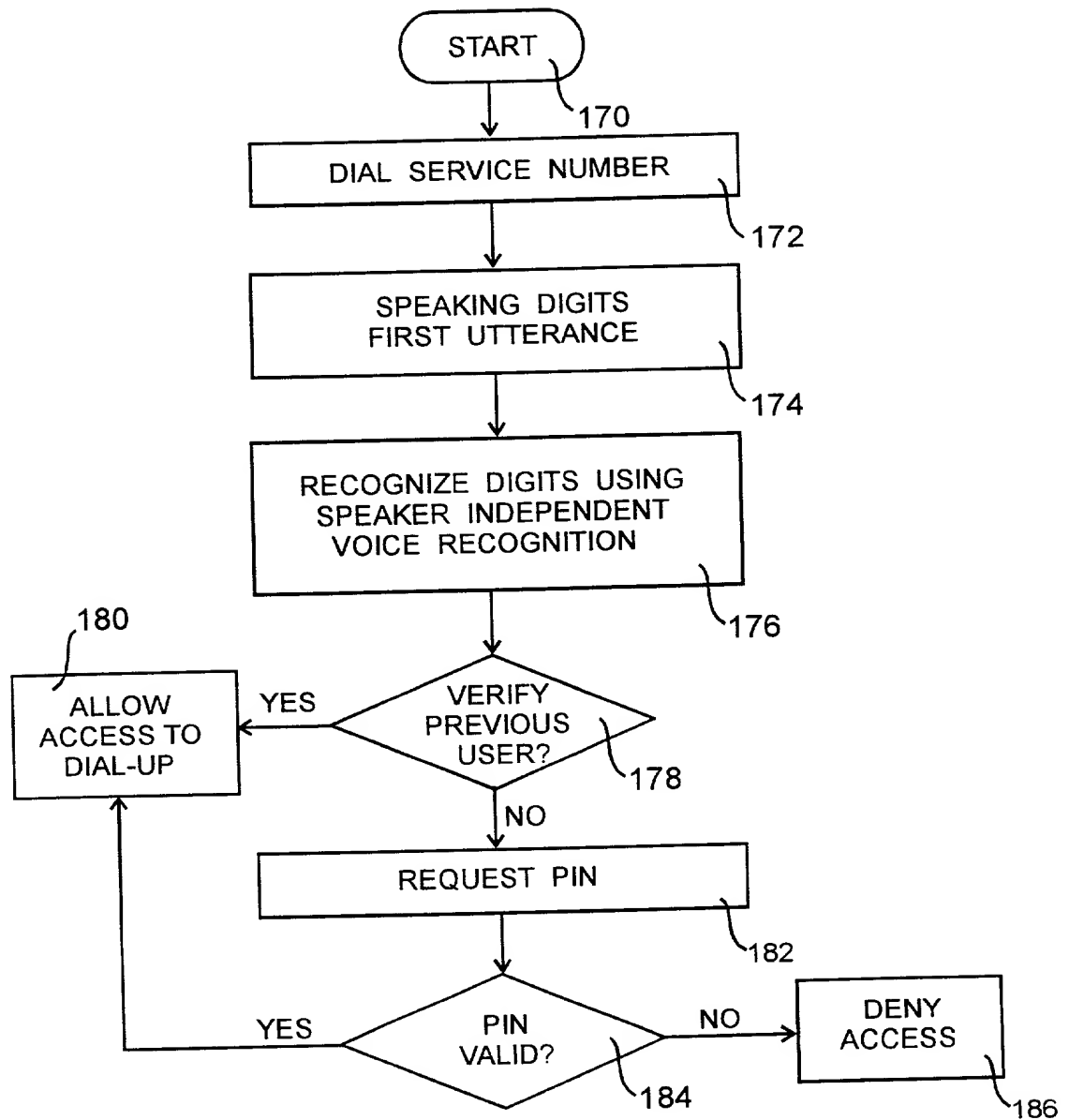
*Fig. 2*

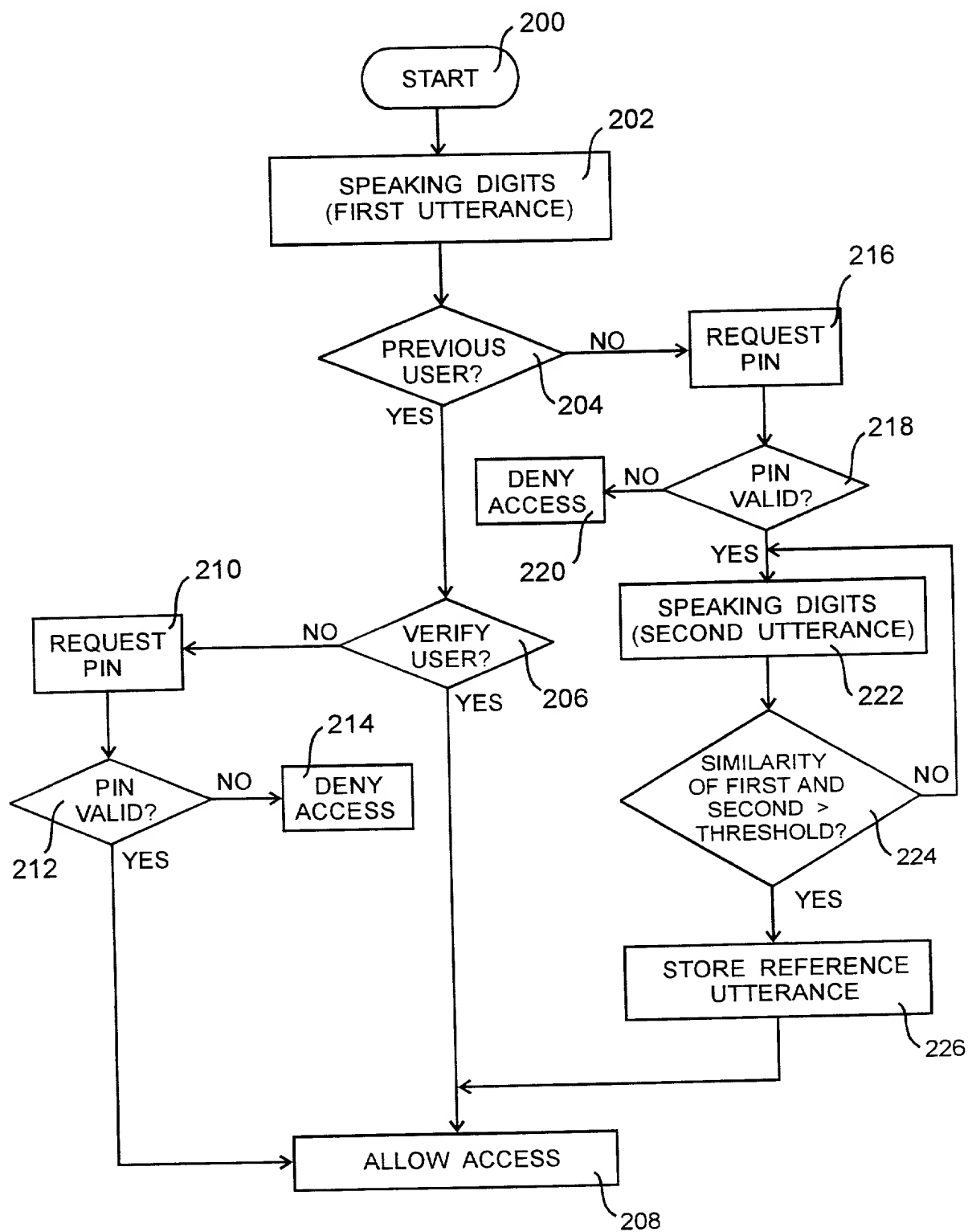
CODE BOOK GENERATION
TRAINING UTTERANCES*Fig. 3*

VERIFICATION DECISION

*Fig. 4*

*Fig. 5*

*Fig. 6*

*Fig. 7*

PATENT APPLICATION DECLARATION COMBINED
WITH POWER OF ATTORNEY

 X REGULAR (UTILITY) OR DESIGN APPLICATION
(check one)

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: "Speaker Verification Method" the specification of which:

 is attached hereto.

 X was filed on May 27, 1997 as U.S. Application Serial No. 08/863,444
and was amended on (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulation, Section 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s):

 X no such applications filed

 such applications identified as follows:

Priority Claimed

| (Serial No.) | (Country) | (Day/Month/Year Filed) | Yes | No |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| <u> </u> | <u> </u> | <u> </u> | <u> </u> | <u> </u> |
| <u> </u> | <u> </u> | <u> </u> | <u> </u> | <u> </u> |
| <u> </u> | <u> </u> | <u> </u> | <u> </u> | <u> </u> |

I hereby claim the priority benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which is material to the examination of this application and which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Prior U.S. Application(s):

(Check
one)X no such applications filed such applications identified as follows:

| (Application Serial No.) | (Filing Date) | (Status) | (Patented, Pending, Abandoned) |
|--------------------------|---------------|----------|--------------------------------|
| | | | |
| | | | |
| | | | |

| (Application Serial No.) | (Filing Date) | (Status) | (Patented, Pending, Abandoned) |
|--------------------------|---------------|----------|--------------------------------|
| | | | |
| | | | |
| | | | |

| (Application Serial No.) | (Filing Date) | (Status) | (Patented, Pending, Abandoned) |
|--------------------------|---------------|----------|--------------------------------|
| | | | |
| | | | |
| | | | |

I hereby declare that: as to any claimed subject matter of this application which is common to my earlier United States or foreign application(s), if any, which I have identified above and claimed the benefit of priority thereof, I do not believe that the same was ever known or used in the United States before my invention thereof or patented or described in any printed publication in any country before my invention thereof or more than one year prior to the first of said earlier application(s), or in public use or on sale in the United States more than one year prior to the first of said earlier application(s), and that the common subject matter has not been patented or made the subject of an inventor's certificate before the date of the first of said earlier U.S. application(s) in any country foreign to the United States on an application, filed by me or my legal representatives or assigns more than twelve months (six months if the present application is a Design patent application) prior to the first of said earlier U.S. application(s), if any; and that, as to any claimed subject matter of this application which is not common to said earlier application(s), if any, I do not know and do not believe that the same was ever known or used in the United States before my invention thereof or patented or described in any printed publication in any country before my invention thereof or more than one year prior to the date of this application, or in public use or on sale in the United States more than one year prior to the date of this application, and that said subject matter has not been patented or made the subject of an inventor's certificate in any country foreign to the United States on an application filed by me or my legal representatives or assigns more than twelve months (six months if the present application is a design patent application) prior to the date of this application.

I HEREBY APPOINT THE FOLLOWING AS MY ATTORNEY(S) OR AGENT(S) WITH FULL POWER OF SUBSTITUTION TO PROSECUTE THIS APPLICATION AND TRANSACT ALL BUSINESS IN THE PATENT AND TRADEMARK OFFICE CONNECTED THEREWITH:

| NAMES | REG. NO. |
|-------------------|----------|
| Bruce E. Stuckman | 36,693 |
| Dale B. Halling | 38,170 |

Send Written Correspondence to: Bruce E. Stuckman
 4H80
 Ameritech Corporation
 2000 W. Ameritech Center Drive
 Hoffman Estates, IL 60196-1025

Direct Telephone Calls to: Bruce E. Stuckman at 847/248-8816

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statement and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

| | | | |
|-----------------------------|---|--------------------------|----------|
| Full name of first inventor | <u>Robert Wesley Bossemeyer, Jr.</u> | | |
| Inventor's signature | <u><i>Robert Wesley Bossemeyer, Jr.</i></u> | | |
| Date: | <u>July 22, 1997</u> | | |
| Residence: | <u>St. Charles, Illinois</u> | <u></u> | |
| | City | State or Foreign Country | |
| Citizenship | <u>United States of America</u> | | |
| Post Office Address | <u>41W091 Colson Drive</u> | | |
| | <u>St. Charles, IL 60175</u> | <u></u> | <u></u> |
| | City | State or Country | Zip Code |

| | | | |
|------------------------------|------------------------------------|--------------------------|----------|
| Full name of second inventor | <u>Rapeepat Ratasuk</u> | | |
| Inventor's signature | <u><i>Rapeepat Ratasuk</i></u> | | |
| Date: | <u>July 29, 1997</u> | | |
| Residence: | <u>Evanston, Illinois</u> | <u></u> | |
| | City | State or Foreign Country | |
| Citizenship | <u>Thailand</u> | | |
| Post Office Address | <u>860 Hinman Avenue, Apt. 616</u> | | |
| | <u>Evanston, IL 60202</u> | <u></u> | <u></u> |
| | City | State or Country | Zip Code |